On the local and global convergence of a reduced quasi-Newton method

Jean Charles Gilbert

INRIA

CENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
BP 105
78153 Le Chesnay Cedex
France
Tél. (1) 39 63 55 11

Rapports de Recherche

N° 565

ON THE LOCAL AND GLOBAL CONVERGENCE OF A REDUCED QUASI-NEWTON METHOD

Jean Charles GILBERT

Septembre 1986

August 1986

ON THE LOCAL AND GLOBAL CONVERGENCE OF A REDUCED QUASI-NEWTON METHOD

SUR LA CONVERGENCE LOCALE ET GLOBALE D' UNE METHODE DE QUASI-NEWTON REDUITE

Jean Charles GILBERT

INRIA

Domaine de Voluceau, Rocquencourt, B.P.105

78153 Le Chesnay Cedex (France)

ON THE LOCAL AND GLOBAL CONVERGENCE OF A REDUCED QUASI-NEWTON METHOD

SUR LA CONVERGENCE LOCALE ET GLOBALE D'UNE METHODE DE QUASI-NEWTON REDUITE

Jean Charles GILBERT

Abstract : In optimization in $\mathbb{R}^n$ with m nonlinear equality constraints, we study the local convergence of reduced quasi-Newton methods that are methods in which the updated matrix is of order n-m. In particular, we give conditions for q-superlinear convergence (in one step). We introduce a device to globalize the local algorithm which consists in determining a step on an arc in order to decrease an exact penalty function. We give conditions so that eventually the step will be equal to one.

Résumé : En optimisation dans $\mathbb{R}^n$ avec m contraintes d'égalité non linéaires, on étudie la convergence locale des méthodes de quasi-Newton réduites qui sont des méthodes où la matrice à mettre à jour est d'ordre n-m. En particulier, on établit les conditions de convergence q-superlinéaire (en un pas). On introduit également une méthode de globalisation de l'algorithme local consistant à déterminer un pas sur un arc de courbe de manière à faire décroître une fonctionnelle pénalisée exacte. On donne des conditions assurant l'admissibilité asymptotique du pas unité.

Abbreviated title : A reduced quasi-Newton method.

Key words : Constrained Optimization, Successive Quadratic Programming, Reduced Quasi-Newton Method, Superlinear Convergence, Exact Penalty Function, Arc Search, Step-size Selection Procedure, Global Convergence.

Subject classification AMS(MOS) : 49D05, 49D30, 65K05.

1 - Introduction

Let ω be an open convex set in $\mathbb{R}^n$ and $f : \omega \longrightarrow \mathbb{R}$ and $c : \omega \longrightarrow \mathbb{R}^m$ $(m < n)$ be two functions of class C_b^ν with $\nu \geq 3$, that is to say that f and c are supposed to be three times continuously differentiable with bounded derivatives on ω. We shall endow $\mathbb{R}^n$ with its canonical basis. We are interested in finding an algorithm for solving the following minimization problem with equality constraints :

$$(1.1) \qquad \min \{ f(x) : x \in \omega , c(x) = 0 \}.$$

In addition to the smoothness of f and c, we shall assume that c is a submersion on ω, that is to say that the m x n jacobian matrix

$$(1.2) \qquad A_x := A(x)$$

of partial derivatives is supposed to be surjective for all x in ω. If ω is "large", this is a very strong hypothesis but it is usual to suppose that the gradients of the constraints are linearly independant at a solution of (1.1) and then, this hypothesis is satisfied in a neighbourhood of a solution. Then, if x_* is a local minimizer for problem (1.1), there exists a unique Lagrange multiplier λ_* so that the first order optimality conditions are satisfied :

$$(1.3) \qquad \left| \begin{array}{l} c(x_*) = 0 \\ \nabla f(x_*) + A_*^T \lambda_* = 0 \end{array} \right.$$

where $\nabla f(x_*)$ is the vector of partial derivatives of f at x_* and $A_* := A(x_*)$. The second equation is the first derivative with respect to x of the lagrangian $l(x,\lambda) := f(x) + (\lambda, c(x))$ at (x_*, λ_*). The second order sufficient condition will also be assumed : the n x n hessian matrix L_* of second derivatives with respect to x of l at (x_*, λ_*) is supposed to be positive definite in the tangent space $N(A_*)$, the kernel of A_*, to the constraints at x_*. For further references, we gather those hypotheses under the name of

assumption A :

- f, c are in $C_b^\nu(\omega)$ with $\nu \geq 3$,

- c is a submersion,
- (x_*, λ_*) satisfies (1.3),
- $h^T L_* h > 0$ for all h in $\mathbb{R}^n$ with $h \neq 0$ and $A_* h = 0$.

Quasi-Newton methods also called variable metric or secant methods are methods for solving a system of nonlinear equations on $\mathbb{R}^N$, say $F(x_*) = 0$, that generate a sequence of points (x_k) and a sequence of matrices (J_k) of order N from the data of a point x_1 and a matrix J_1 by the formula :

$$x_{k+1} = x_k - J_k^{-1} F(x_k),$$

where J_k is supposed to be nonsingular and updated at each iteration by the following scheme :

$$J_{k+1} := U(J_k, \gamma_k, \sigma_k),$$
$$\gamma_k := F(x_{k+1}) - F(x_k),$$
$$\sigma_k := x_{k+1} - x_k.$$

The rule U is designed in order that J_{k+1} will satisfy the secant equation $J_{k+1}\sigma_k = \gamma_k$ and then improves the approximation by J_k of the jacobian matrix $F'(x_*)$ at the solution x_*. These methods are particularly attractive because second order derivatives need not be calculated and that a q-superlinear rate of convergence for (x_k) can be obtained (see the review paper of Dennis-Moré (1977)), that is to say :

$$\frac{||x_{k+1} - x_*||}{||x_k - x_*||} \longrightarrow 0 \quad \text{as} \quad k \longrightarrow \infty. \tag{1.4}$$

Such a method can be used for solving system (1.3), but matrices of order n+m will have to be updated. The aim of this paper is to introduce and study such quasi-Newton methods but with updated matrices of order n-m.

The Successive Quadratic Programming (SQP) proposed by Wilson (1963) and Han (1976) improves the previous method with regard to the order of the updated

matrices. In this method x_{k+1} is obtained from x_k by solving a quadratic program with linear constraints :

$$(1.5) \qquad \left| \begin{array}{l} \min \ [f'(x_k).(x - x_k) + \frac{1}{2} (x - x_k)^T L_k (x - x_k)] \\ x \in \mathbb{R}^n \ \text{ and } \ c(x_k) + c'(x_k).(x - x_k) = 0 \end{array} \right.$$

where we have used a dot to separate the linear operators $f'(x_k)$ and $c'(x_k)$ from their argument $(x-x_k)$. The matrix L_k of order n is updated in order to approximate L_*. The fact that the lagrangian has a minimum at x_* in the tangent plane to the constraints at x_* allows to understand the method : it minimizes a quadratic approximation of the lagrangian subject to the linearized constraints. Under assumption A, the solution of (1.5) can be written in the form (see Gabay (1982,b)) :

$$(1.6) \qquad x_{k+1} = x_k - A_k^- c(x_k) - Z_k^- (Z_k^{-T} L_k Z_k^-)^{-1} [g(x_k) - Z_k^{-T} L_k A_k^- c(x_k)],$$

where A_k^- is any right inverse of $A_k := A(x_k)$, Z_k^- is any $n \times (n-m)$ matrix whose columns form a basis of the tangent space $N(A_k)$ to the manifold $M_k := c^{-1}(c(x_k))$ at x_k and $g(x_k)$ is the reduced gradient defined by

$$(1.7) \qquad g(x_k) := Z_k^{-T} \nabla f(x_k) \ \in \mathbb{R}^{n-m}.$$

The first part $(-A_k^- c(x_k))$ of the displacement in (1.6) is a restoration step, i.e. a Newton-like step for solving $c(x_*) = 0$. It belongs to $R(A_k^-)$ which is a complementary space to $N(A_k)$ in $\mathbb{R}^n$. The second part of the displacement in (1.6) is a minimization step belonging to the tangent space $R(Z_k^-) = N(A_k)$ to M_k at x_k.

Let us consider for a while the case where the constraints are linear :

$$(1.8) \qquad c(x) := Ax + b = 0,$$

where A is a $m \times n$ surjective matrix and b is a vector in $\mathbb{R}^m$. Like in the nonlinear case, let us introduce Z^-, a $n \times (n-m)$ matrix whose columns generate $N(A)$: $AZ^- = 0$. Suppose that the first iterate belongs to the plane of the constraints : $c(x_1) = 0$. Then, any point x satisfying the constraints (1.8) can be expressed by using a reduced variable u in $\mathbb{R}^{n-m}$ $(x = x_1 + Z^- u)$ and the

problem of minimizing f subject to the constraints (1.8) is equivalent to the one of minimizing $\varphi(u) := f(x_1+Z^- u)$ on $\{ u \in \mathbb{R}^{n-m} : x_1 + Z^- u \in \omega \}$:

$$\min \{ \varphi(u) : u \in \mathbb{R}^{n-m} , x_1 + Z^- u \in \omega \}. \tag{1.9}$$

By considering the optimality equation $\nabla\varphi(u_*) = Z^{-T} \nabla f(x_1+Z^- u_*) = 0$, a quasi-Newton method for solving problem (1.9) generates a sequence (u_k) in $\mathbb{R}^{n-m}$ and a sequence (G_k) of nonsingular matrices of order n-m such that

$$u_{k+1} = u_k - G_k^{-1} \nabla\varphi(u_k).$$

By setting $x_k := x_1 + Z^- u_k$, we obtain

$$x_{k+1} = x_k - Z^- G_k^{-1} g(x_k), \tag{1.10}$$

where $g(x_k) := Z^{-T} \nabla f(x_k)$ is the reduced gradient of f at x_k and G_k is updated as follows :

$$\begin{aligned} G_{k+1} &= U(G_k, \gamma_k, \sigma_k), \\ \gamma_k &= g(x_{k+1}) - g(x_k), \\ Z^- \sigma_k &= x_{k+1} - x_k, \end{aligned} \tag{1.11}$$

in order to approximate $\nabla^2\varphi(u_*)$, the hessian of φ at u_*, which is also the <u>reduced hessian of f at</u> x_* :

$$Z^{-T} \nabla^2 f(x_*) Z^-.$$

The conditions so that the rate of convergence of the sequence (u_k) will be q-superlinear can be satisfied and because of the injectivity of Z^-, the same conditions will assure the q-superlinear convergence of the sequence (x_k).

The algorithm (1.10)-(1.11) is a <u>reduced quasi-Newton method</u> because the order of the updated matrices is n-m rather than n in the SQP. Our aim is to study such methods in case of nonlinear constraints. They are particularly well adapted to problems with n large and n-m small. That situation appears for example in the parametric identification of nonlinear sources in partial differential equations. If a finite element method is used to discretize the state equations (the constraints), m is large, say of the order of 1000, whereas the number n-m

of identifiable parameters is usually small : 2 or 3 in the example of Blum-Gilbert-Thooris (1985). In that case, a reduced quasi-Newton method is usable while the SQP is not because of the order of the matrix that should be updated. Another advantage in developping reduced quasi-Newton methods comes from the fact that, under assumption A, the projected hessian of the lagrangian is positive definite at the optimum. Therefore, positive definite quasi-Newton approximations of that operator can be generated, in particular by the BFGS update formula which possesses some stability properties. We see that reduced quasi-Newton methods appears rather natural. So, it is important to generalize the algorithm (1.10)-(1.11) in case the constraints are nonlinear.

This can be done by using the implicit function theorem in order to obtain a reduced objective function :

$$\varphi(u) = f(\xi(u)),$$

where $\xi : u \in V \subset \mathbb{R}^{n-m} \longrightarrow \xi(u) \in \omega \subset \mathbb{R}^n$ is a parametric representation of the regular manifold $c^{-1}(0)$ around $x_* := \xi(u_*) : c(\xi(u)) = 0$ for all u in V. This is the basic idea of methods like the generalized reduced gradient (GRG) method (Abadie-Carpentier (1969)) or the gradient projection method (Rosen (1961)). In fact, the parametric representation $\xi(u)$ is usually not known and this leads to several difficulties. Because the method asks the generated sequence (x_k) to be feasible ($c(x_k) = 0$ for all k) and that this cannot be achieved exactly in practice, some criterion has to be introduced to decide when to stop the restauration steps, i.e. how well the equality $x_{k+1} = \xi(u_{k+1})$ has to be realized (Mukai-Polak (1978)). Another difficulty appears when x_k is far from x_* and that a step-size has to be introduced in the u-space in order to globalize the method. Indeed, everytime a step-size is tried, an infinite number of restoration steps have to be done : see Gabay (1975), Gabay-Luenberger (1976), Mukai-Polak (1978), Gabay (1982,a).

On the other hand, some non-feasible reduced quasi-Newton methods have been developped recently. Gabay (1982,b) has studied the following algorithm :

$$x_{k+1} = x_k + r_k^1 + t_k^1, \tag{1.12}$$

$$r_k^1 := - A_k^- c(x_k), \tag{1.13}$$

(1.14) $$t_k^1 := - Z_k^- G_k^{-1} g(x_k),$$

where A_k^- is any right inverse of A_k, Z_k^- is any n x (n-m) matrix whose columns form a basis of the tangent space $N(A_k)$, G_k is a nonsingular matrix of order (n-m) and $g(x_k)$ is the reduced gradient of f at x_k. The tangent step t_k^1 in (1.12), tangent to the manifold M_k, has the same structure as the displacement in (1.10) except for the basis Z_k^- which changes here at each iteration. The restoration step r_k^1 in (1.12) is introduced to improve the feasibility of the sequence. The displacement in (1.12) can also be deduced from the displacement (1.6) of the SQP method by dropping the last part of the minimization step and by considering G_k as an approximation of the projected hessian $Z_k^{-T} L(x_k,\lambda_k) Z_k^-$.

For their part, Coleman and Conn (1982,a) have studied the following algorithm :

(1.15) $$x_{k+1} = x_k + r_k^2 + t_k^2,$$

(1.16) $$r_k^2 := - A_k^- c(x_k+t_k^2),$$

(1.17) $$t_k^2 := - Z_k^- G_k^{-1} g(x_k),$$

where Z_k^- is a n x (n-m) matrix whose columns form an orthogonal basis of $N(A_k)$ and A_k^- is the Penrose pseudo-inverse of A_k : $A_k^- := A_k^T (A_k A_k^T)^{-1}$. In fact,the orthogonality of the columns of Z_k^- is not essential and any right inverse for A_k can be used. The only relevant difference with the algorithm of Gabay lies in the restoration step in which the constraints are evaluated at $x_k+t_k^2$, after the tangent step, rather than at x_k in the algorithm (1.12)-(1.14).

The study of both algorithms (1.12)-(1.14) and (1.15)-(1.17) shows that when the matrices G_k are suitably chosen and the initial point x_1 is close to x_*, the sequence (x_k) generated by any of those algorithms converges to x_* q-superlinearly in two steps, that is to say :

$$\frac{||x_{k+1} - x_*||}{||x_{k-1} - x_*||} \longrightarrow 0 \quad \text{as} \quad k \longrightarrow \infty.$$

This rate of convergence is not so good as than the rate (1.4) obtained with reduced quasi-Newton methods when the constraints are linear. On the other hand, counter-examples have been given by Byrd (1985) and Yuan (1985) for which

both methods of Gabay and Coleman-Conn do not behave better than with a two step q-superlinear rate of convergence. Therefore, we may ask the question : is this fact the price to pay for the generalization of reduced quasi-Newton methods to nonlinear constraints ?

The question is considered in the subsequent sections. If unconstrained optimization problems are closely related to the solution of nonlinear equations (via the optimality condition $f'(x_*) = 0$), we see from (1.3) that constrained optimization problems are related to the solution of two coupled nonlinear equations :

$$c(x_*) = 0 \quad \text{in } \mathbb{R}^m \tag{1.18}$$

$$g(x_*) = 0 \quad \text{in } \mathbb{R}^{n-m} \tag{1.19}$$

The equation (1.19) expresses the vanishment of the reduced gradient defined in (1.7) and is obtained by projecting the second equation of (1.3*) on the tangent space $N(A_*)$. A "decoupling" method for solving (1.18)-(1.19) is introduced in section 2. At x_k, the first part of the step of the method consists in doing a Newton-like displacement for solving (1.18). This leads to a point y_k. Then, x_{k+1} is obtained by doing a Newton-like displacement for solving (1.19) from the point y_k :

$$y_k = x_k - \bar{A}_* c(x_k), \tag{1.20}$$

$$x_{k+1} = y_k - \bar{B}_* g(y_k), \tag{1.21}$$

where $\bar{A}_*$ is a right inverse of $\nabla c(x_*)$ and $\bar{B}_*$ is a right inverse of $\nabla g(x_*)$. We shall see that only $\bar{B}_*$ has to be imposed in order to ensure local q-quadratic convergence (in one step) of the process. In section 4, we apply this algorithm to constrained optimization, when g in (1.19) has the special structure (1.7), and it takes the form of a reduced Newton method. Its extension to reduced quasi-Newton method is then easily done. In section 5, necessary and sufficient conditions that assure the q-superlinear convergence (in one step) of the method are established.

In fact the obtained algorithm appears to be Coleman-Conn's algorithm but the considered sequence is their sequence $(x_k+t_k^2)$ rather than their sequence (x_k).

So, our result of this sections consists, on the one hand, in making a link between the algorithm of Coleman-Conn for constrained optimization problems and algorithms for solving two coupled equations, which gives some insight into the method and, on the other hand, in obtaining necessary and sufficient conditions for q-superlinear convergence in one step of the sequence (x_k) whereas the result of Coleman-Conn (1982,a) concerns the q-superlinear convergence in two steps of the sequence (y_k) of our method.

The globalization of the local method could then be done like in the paper of Coleman and Conn (1982,b). In section 6, however, we examine another globalizing technique essentially based on the ideas of Han (1977) for the SQP (see also Danilin-Pschenichny (1965)). We introduced the following exact penalty function :

$$\Theta_p(x) = f(x) + p\,||c(x)||_1, \tag{1.22}$$

where p is a large enough penalty parameter and $||.||_1$ is the 1-norm on $\mathbb{R}^m$. We look for x_* by minimizing Θ_p on ω. The idea is then to obtain a descent direction for Θ_p at the current iterate from the displacement calculated by the local algorithm (1.20)-(1.21). Contrary to what happens with the SQP, our total displacement is not a descent direction for Θ_p any more. So we shall introduce a descent arc, being inspired in that way by the work of Gabay (1982,b) for his algorithm and Mayne-Polak (1982) for the SQP, although in those algorithms, the arc was introduced for other reasons. A search on the arc is done in order to decrease the penalty function Θ_p with the help of an Armijo-like criterion. This gives a theorem assuring the global convergence of the method. An advantage of that technique is that under natural conditions the step-size is equal to one after a finite number of iterations. Therefore there is a sweet transition from the global to the local method that does not prevent the q-superlinear convergence to occur.

Just a word on the notations. If (v_k) is a sequence in a normed space $(E,||.||_E)$ and (α_k) is a sequence of positive numbers, we shall say that (v_k) is a large O of (α_k) (we shall note $v_k = O(\alpha_k)$) if the sequence $(||v_k||_E/\alpha_k)$ is bounded and we shall say that (v_k) is a small o of (α_k) (we shall note $v_k = o(\alpha_k)$) if the sequence $(||v_k||_E/\alpha_k)$ converges to zero. We shall note v^i, the i-th component

of a vector v in E. If A is a linear operator from $(E,||.||_E)$ to $(F,||.||_F)$, we shall note $||A|| = \sup \{ ||Av||_F : ||v||_E \leq 1 \}$.

This paper constitutes a revised version of a part of the report number RR-482 of INRIA in which some techniques for updating the reduced matrix have also been investigated. A variant of the method is given in Gilbert (1986,b).

2 - A decoupling method for solving two nonlinear coupled equations

Let us consider the following coupled system of nonlinear equations :

$$\left| \begin{array}{l} F(x) = 0 \\ \\ G(x) = 0 \end{array} \right. \tag{2.1}$$

where $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $(m < n)$ and $G : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-m}$ are supposed to be smooth. Let x_* be a solution of (2.1) and let us note A_* the m x n jacobian matrix of F at x_* and B_* the (n-m) x n jacobian matrix of G at x_*. We shall say that x_* is a regular solution of (2.1) if the jacobian matrix of the system (2.1),

$$J_* := \begin{bmatrix} A_* \\ B_* \end{bmatrix}, \tag{2.2}$$

is nonsingular. This will be the case if and only if $N(A_*) \wedge N(B_*) = \{0\}$. We would like to define a Newton-like method for solving (2.1) without having to inverse the jacobian $J(x_k)$ of (2.1) at x_k. For that, let us suppose that x_* is a regular solution. Then A_* and B_* are surjective and we can introduce a right inverse A_*^- of A_* and a right inverse B_*^- of B_* :

$$A_* A_*^- = I_m , \quad B_* B_*^- = I_{n-m}. \tag{2.3}$$

Two algorithms using A_*^- and B_*^- can be considered and we defined them with the help of the fixed point maps ξ_1 and ξ_2. The first one is

$$x_{k+1} = \xi_1(x_k), \tag{2.4}$$

$$\xi_1(x) := x - A_*^- F(x) - B_*^- G(x) \tag{2.5}$$

and the second one is

$$x_{k+1} = \xi_2(x_k) := (\psi \circ \varphi)(x_k), \tag{2.6}$$

$$\varphi(x) := x - A_*^- F(x), \tag{2.7}$$

$$\psi(y) := y - B_*^- G(y). \tag{2.8}$$

Those algorithms are somewhat "ideal". Indeed, the matrices A_* and B_* are not known and so neither are A_*^- and B_*^-. But not better than a good "ideal" Newton method to introduce quasi-Newton methods. The relations (2.3) do not determined the right inverses A_*^- and B_*^- completely. Therefore, we can try to choose them so that the sequences generated by the algorithms (2.4)-(2.5) and (2.6)-(2.8) will have a good local behavior. The next two propositions show that this is possible : we can get conditions on A_*^- and B_*^- in order to have $\xi_1'(x_*) = 0$ and $\xi_2'(x_*) = 0$ which will ensure the q-quadratic rate of convergence for both algorithms. We shall say that a n row matrix is a basis of a given subspace of $\mathbb{R}^n$ if its columns form a basis of that subspace.

Proposition 2.1 : Suppose that F and G are differentiable at x_*, a regular solution of (2.1). Then, the following statements are equivalent :

(i) $\xi_1'(x_*) = 0$,

(ii) $R(A_*^-) = N(B_*)$ and $R(B_*^-) = N(A_*)$,

(iii) for any right inverse $A_*^\sim$ of A_* and any basis Z_*^- of $N(A_*)$, we have

$$A_*^- = [I - Z_*^- (B_* Z_*^-)^{-1} B_*] A_*^\sim, \tag{2.9}$$

$$B_*^- = Z_*^- (B_* Z_*^-)^{-1}. \tag{2.10}$$

Proof. First, we prove (i) <=> (ii). Statement (i) is equivalent to

$$I = A_*^- A_* + B_*^- B_*. \tag{2.11}$$

The right hand side of (2.11) is equal to $[A_*^- B_*^-]J_*$. Then (2.11) means that $[A_*^- B_*^-]$ is the inverse of J_* and therefore is equivalent to $I = J_*[A_*^- B_*^-]$, that is to say that $A_*B_*^- = 0$ and $B_*A_*^- = 0$, which is statement (ii) because A_*^- and B_*^- are injective. Next, we prove (ii) => (iii). Let Z_*^- be any basis of $N(A_*)$: $R(Z_*^-) = N(A_*)$. Because J_* is nonsingular, $B_*Z_*^-$ is nonsingular. By multiplying (2.11) to the right by Z_*^-, we get $Z_*^- = B_*^-(B_*Z_*^-)$ and therefore (2.10). (2.9) is obtained by multiplying (2.11) to the right by any right inverse $A_*^\sim$ of A_* and by

using (2.10). It remains to prove (iii) => (ii). If we take $\tilde{A}_*^- = A_*^-$ in (2.9), we obtain $B_* A_*^- = 0$ and because $A_* Z_*^- = 0$, (2.10) gives $A_* B_*^- = 0$ •

Proposition 2.2 : Suppose that F and G are differentiable at x_*, a regular solution of (2.1). Then, the following statements are equivalent :

(i) $\xi_2'(x_*) = 0$,
(ii) $R(B_*^-) = N(A_*)$,
(iii) for any basis Z_*^- of $N(A_*)$, we have $B_*^- = Z_*^- (B_* Z_*^-)^{-1}$.

Proof. The equivalence (i) <=> (ii) comes from the fact that $\xi_2'(x_*) = (I - B_*^- B_*)(I - A_*^- A_*)$ and that the spaces $N(A_*) = R(I - A_*^- A_*)$ and $R(B_*^-) = N(I - B_*^- B_*)$ have the same dimension n-m. To prove (iii) from (i), let Z_*^- be any basis of $N(A_*)$. By multiplying

$$(I - B_*^- B_*)(I - A_*^- A_*) = 0$$

to the right by Z_*^-, we get $Z_*^- = B_*^-(B_* Z_*^-)$ and therefore (iii) because $B_* Z_*^-$ is nonsingular. From (iii), we get (ii) by multiplying to the left $B_*^- = Z_*^-(B_* Z_*^-)^{-1}$ by A_* •

In the statement (iii) of proposition 2.1, we could equivalently have given to B_* the role of A_*. We also see that the right inverses A_*^- and B_*^- are completely determined by the condition (i) of proposition 2.1 and do not depend on the choice of $\tilde{A}_*^-$ and Z_*^- in (iii). Similarly, the right inverse B_*^- is completely determined by condition (i) of proposition 2.2 and does not depend on the choice of Z_*^- in (iii).

From proposition 2.1, we see that $\xi_1'(x_*) = 0$ if and only if $[A_*^- \; B_*^-]$ is the inverse of J_*. This means that the algorithm (2.4)-(2.5) is in fact the "ideal" (with J_*^{-1} rather than $J(x_k)^{-1}$) Newton method for solving (2.1) (see the displacement in (2.5)), the method we wanted to avoid. On the other hand, proposition 2.2 shows that the algorithm (2.6)-(2.8) need less conditions to have a good local behaviour than the algorithm (2.4)-(2.5). We shall call it the ideal decoupling method. The fact that no conditions are required on the right inverse A_*^- means that any solver of the first equation in (2.1) can be used in (2.7),

independently of the second equation of (2.1), whereas this is not true for the solver B_*^- of the second equation of (2.1) that has to be adapted to the first equation.

The results of proposition 2.1 and 2.2 have a geometrical interpretation. In the ideal Newton method, (x_k) will converge q-quadratically if the displacements $(-A_*^- F(x_k))$ and $(-B_*^- G(x_k))$ belong to the tangent space at x_* to the manifolds defined respectively by the pre-image of 0 by F and G. In the ideal decoupling method, only the second step $(-B_*^- G(y_k))$ has to belong to the tangent space $N(A_*)$, the first step is arbitrary.

3 - A change of coordinates

Before applying the results of the previous section to constrained optimization, let us give some examples of right inverses A_x^- of A_x and basis Z_x^- of $N(A_x)$. For more details, we refer to Gabay (1982,a).

Once the injective matrices A_x^- and Z_x^- have been chosen, the columns of $[A_x^- \; Z_x^-]$ form a new basis of $\mathbb{R}^n$. Indeed, $R(A_x^-)$ is a complementary space of $N(A_x)$. To make a change of coordinates in that new basis, it is convenient to introduce the additional (n-m) x n matrix Z_x given in the following proposition.

Proposition 3.1 : Let A_x be a m x n (m < n) surjective matrix, A_x^- be any right inverse of A_x and Z_x^- be any basis of $N(A_x)$. Then, there exists a unique (n-m) x n matrix Z_x such that

$$Z_x A_x^- = 0, \tag{3.1}$$

$$Z_x Z_x^- = I_{n-m}. \tag{3.2}$$

Furthermore, we have

$$I_n = A_x^- A_x + Z_x^- Z_x. \tag{3.3}$$

Proof. Existence and unicity of the matrix Z_x come from the nonsingularity of $[A^- \; Z^-]$ and (3.3) comes from the fact that $[A^T \; Z_x^T]^T$ is the inverse of $[A^- \; Z^-]$ •

The relation (3.1) shows that $N(Z_x) = R(A_x^-)$ and (3.2) shows that Z_x^- is a right inverse of Z_x. The exponent (-) always means that a linear problem has to be solved when we apply the matrix to some vector. The equality (3.3) can be used to introduce a change of coordinates. Indeed, by applying it to a vector ξ of $\mathbb{R}^n$, we see that $A_x\xi$ are the coordinates of ξ in $R(A_x^-) = N(Z_x)$ and $Z_x\xi$ are the coordinates of ξ in $R(Z_x^-) = N(A_x)$.

The first choice of matrices A_x^- and Z_x^- is frequently made in constrained optimization and could be called the orthogonal framework : A_x^- is the Moore-Penrose pseudo-inverse of A_x (see Ben-Israel-Greville (1974)) and Z_x^- is an orthogonal basis of $N(A_x)$ for the scalar product $(\xi,\eta) = \sum_{i=1}^{n} \xi^i \eta^i$. We have :

$$
\begin{aligned}
&A_x^- = A_x^T (A_x A_x^T)^{-1}, \\
&Z_x^{-T} Z_x^- = I_{n-m}.
\end{aligned}
\tag{3.4}
$$

Then $Z_x = Z_x^{-T}$ is the unique matrix satisfying (3.1) and (3.2). We see that $R(A_x^-)$ is orthogonal to $N(A_x)$.

Another choice of matrices A_x^- and Z_x^- is made when a separation of the variables occurs naturally, as in optimal control problems or in parameters identification problems. This is also the framework adopted to introduce the GRG method (Abadie-Carpentier (1969)) and could be called the partitioned framework : A_x is supposed to be partitioned in two submatrices

$$A_x = [\, C_x \quad D_x \,],$$

where the m x m matrix C_x is nonsingular and D_x has dimension m x (n-m). The right inverse A_x^- is then taken as

$$A_x^- = \begin{bmatrix} C_x^{-1} \\ 0 \end{bmatrix}$$

and the basis of $N(A_x)$ is

$$Z_x^- = \begin{bmatrix} - C_x^{-1} D_x \\ I_{n-m} \end{bmatrix}.$$

Then $Z_x = [0\ I_{n-m}]$ is the unique matrix satisfying (3.1) and (3.2).

In the following, we shall suppose that the choice of $(\bar{A}_x, \bar{Z}_x)$ is a smooth function of x :

assumption B :

the function $x \longrightarrow (\bar{A}_x, \bar{Z}_x, Z_x)$ is bounded on ω and is in $C_b^{\nu-1}(\omega)$ with $\nu \geq 3$.

On this regularity hypothesis, see Gill-Murray-Saunders-Stewart-Wright (1985) and Byrd-Schnabel (1986).

4 - A reduced quasi-Newton method for constrained optimization

In this section, we apply the results of section 2 to constrained optimization. The first step consists in reducing the size of the optimality system (1.3). This can be done because the second optimality condition can be expressed by n-m equations rather than n, in fact, by the nullity of the n-m coordinates of the orthogonal projection of $\nabla f(x_*)$ on $N(A_*)$. If $\bar{Z}_*$ is any basis of $N(A_*)$, the orthogonal projector on $N(A_*)$ is $\bar{Z}_*(\bar{Z}_*^T\bar{Z}_*)^{-1}\bar{Z}_*^T$. Then, the second equation of (1.3) is projected on $N(A_*)$ by multiplying it by $\bar{Z}_*^{-T}$. Using the definition (1.7) of the reduced gradient, the system (1.3) can be rewritten as follows :

$$(4.1) \qquad \left| \begin{array}{l} c(x_*) = 0 \\ \\ g(x_*) = 0 \end{array} \right.$$

In order to apply the previous results, we need to calculate the first derivative of g at x_*. This can be done like in Stoer (1984) and Nocedal-Overton (1985) :

$$(4.2) \qquad \nabla g(x_*) = \nabla(\bar{Z}_x^{-T}\ (\nabla f(x) + A_x^T\ \lambda_*))(x_*) = \bar{Z}_*^{-T}\ L_*.$$

The jacobian matrix of (4.1),

$$\begin{bmatrix} A_* \\ Z_*^{-T} L_* \end{bmatrix},$$

is nonsingular because of the surjectivity of A_* and the second order sufficient condition which is equivalent to the nonsingularity of

$$G_* := Z_*^{-T} L_* Z_*^{-}. \tag{4.3}$$

We shall note $H_* := G_*^{-1}$. Now, let us apply algorithm (2.4)-(2.5) to the system (4.1). By using statement (iii) of proposition 2.1 and (4.2), we get the following q-quadratically convergent algorithm :

$$x_{k+1} = x_k - [I - Z_*^{-} H_* Z_*^{-T} L_*] A_*^{-} c(x_k) - Z_*^{-} H_* g(x_k), \tag{4.4}$$

where A_*^{-} is any right inverse of A_* (playing the role of $A_*^{\sim}$ in (2.9)) and Z_*^{-} is any basis of $N(A_*)$. This is exactly the "ideal" SQP method (see (1.6)). So, we obtain the result of Nocedal-Overton (1985) according to which the SQP method is a Newton method for solving (4.1). We shall not go further with that method. If we apply algorithm (2.6)-(2.8) to the system (4.1), we get, by using statement (iii) of proposition 2.2 and (4.2) :

$$\bar{y}_k = x_k - A_*^{-} c(x_k), \tag{4.5}$$

$$\bar{x}_{k+1} = \bar{y}_k - Z_*^{-} G_*^{-1} g(\bar{y}_k), \tag{4.6}$$

where A_*^{-} is any right inverse of A_* and Z_*^{-} is any basis of $N(A_*)$. The following lemma is a consequence of proposition 2.2.

Lemma 4.1 : Suppose that assumptions A and B are satisfied and let x_* be a solution of (1.1). There exists a positive constant C that depends only on f and c such that if x_k, $\bar{y}_k$ given by (4.5) and $\bar{x}_{k+1}$ given by (4.6) are in ω, we have

$$||\bar{x}_{k+1} - x_*|| \leq C \, ||x_k - x_*||^2. \tag{4.7}$$

From the q-quadratically convergent algorithm (4.5)-(4.6), a quasi-Newton method is easily introduced. In (4.6), G_* is replaced by an approximation G_k and Z_*^{-} is replaced by $Z(y_k)^{-}$ which intervenes in the calculation of the reduced gradient $g(y_k)$. If A_*^{-} in (4.5) is replaced by $A(x_k)^{-}$, the constraints will have

to be linearized twice per iteration : at x_k to calculate $A(x_k)$ and at y_k to calculate the basis $Z(y_k)^-$. Since the constraints have to be linearized at y_k to calculate the reduced gradient in (4.6), we avoid two linearizations of the constraints by replacing A_*^- in (4.5) by $A(y_{k-1})^-$. So we obtain the following local algorithm :

$$y_k = x_k - A(y_{k-1})^- c(x_k), \tag{4.8}$$

$$x_{k+1} = y_k - Z(y_k)^- G_k^{-1} g(y_k). \tag{4.9}$$

We shall note r_k the restoration step and t_k the tangent step :

$$r_k := - A(y_{k-1})^- c(x_k), \tag{4.10}$$

$$t_k := - Z(y_k)^- G_k^{-1} g(y_k). \tag{4.11}$$

We shall also use the total displacements

$$d_k := r_k + t_k, \tag{4.12}$$

$$e_k := t_k + r_{k+1}. \tag{4.13}$$

Practically, the algorithm cannot start with (4.8) from a point x_1 without knowing a point y_0. So, we shall suppose in the following that the algorithm start with (4.9) from a point y_0 in ω. This is really the same algorithm as Coleman-Conn's method (1.15)-(1.17) if we change in (4.8)-(4.9) y_k by x_k and x_{k+1} by x_k+t_k. But contrary to the sequence (y_k) in (4.8)-(4.9) that does not usually converge q-superlinearly in one step (see the examples of Byrd (1985) and Yuan (1985)), the sequence (x_k) will have some chance to converge q-superlinearly as expected from the behaviour of the ideal algorithm (4.5)-(4.6).

In fact, it is not essential to reduce the size of the optimality system before applying algorithm (2.6)-(2.8). The same method (4.8)-(4.9) can be obtained when the decoupling method is applied to the optimality conditions (1.3). In this case, $B_* = [L_* \; A_*^T]$ and

$$B_*^- = \begin{bmatrix} Z_*^- G_*^{-1} Z_*^{-T} \\ A_*^{-T} (I - L_* Z_*^- G_*^{-1} Z_*^{-T}) \end{bmatrix},$$

where A_*^- is any right inverse of A_* and Z_*^- is any basis of $N(A_*)$. Furthermore, that way to do gives an iteration scheme for the Lagrange multipliers (λ_k) :

$$\lambda_{k+1} = - A(y_k)^{-T} \nabla f(y_k) + A(y_k)^{-T} L_k Z(y_k)^- G_k^{-1} g(y_k), \tag{4.14}$$

where L_k is an approximation of L_*. This formula simplifies the one obtained when the quasi-Newton method is applied to (1.3) (see Gabay (1982,b)).

The algorithm (4.8)-(4.9) is a reduced quasi-Newton method for (x_k) because the only matrix to update is the approximation G_k of G_*. Unfortunately, it is not the case any more for the sequence (λ_k) generated by (4.14) since L_k intervenes in (4.14). The next theorem gives conditions in order to have the local linear convergence of the algorithm.

Theorem 4.2 : Suppose that assumptions A and B are satisfied. There exists a positive constant C that depends only on f, c and ω such that if r is a real number in $]0,1[$ and if

$$||y_0 - x_*|| \leq c\, r, \tag{4.15}$$

$$||G_k - G_*|| \leq c\, r \quad \text{for all } k, \tag{4.16}$$

then the algorithm (4.8)-(4.9) generates from y_0 a sequence (x_k) in ω that converges q-linearly to x_* and

$$||x_{k+1} - x_*|| \leq r\, ||x_k - x_*|| \tag{4.17}$$

for all k.

Proof. According to assumptions A and B, there exists a positive constant C_0 such that

$$\max (||A(y)|| , ||A(y)^-|| , ||Z(y)^-|| , ||g'(y)||) \leq C_0, \tag{4.18}$$

for all y in ω and with $\bar{y}_k$ calculated from x_k by (4.5),

$$\max (||\bar{y}_k - x_*|| , ||y_k - x_*||) \leq C_0\, ||x_k - x_*||, \tag{4.19}$$

as soon as y_{k-1} and x_k are in ω. This last two inequalities are obtained from (4.5) and (4.8) by developing $c(x_k)$ around x_*. Let ε, η and $\bar{\delta}$ be three fixed positive constants such that

$$B(x_*,\varepsilon) \subset \omega, \tag{4.20}$$

$$||G_*^{-1}|| \leq \eta, \tag{4.21}$$

$$\bar{\delta} < \frac{1}{\eta}, \tag{4.22}$$

$$(1 + C_0)\ \bar{\delta} < \varepsilon, \tag{4.23}$$

where $B(x_*,\varepsilon)$ denotes the ball of radius ε centred at x_*. We shall note C_i ($i=1,2,\ldots$) any positive constant that depends only on f, c, ω, $\bar{\delta}$, η and C_0. Let us choose δ in $]0,\bar{\delta}]$ and let us prove that if y_{k-1}, x_k and G_k satisfy

$$||y_{k-1} - x_*|| \leq C_0\ \delta, \tag{4.24}$$

$$||x_k - x_*|| \leq \delta, \tag{4.25}$$

$$||G_k - G_*|| \leq \delta, \tag{4.26}$$

then y_k and x_{k+1} are well defined by (4.8)-(4.9) and that there exists a positive constant C_1 such that

$$||y_k - x_*|| \leq C_0\ \delta, \tag{4.27}$$

$$||x_{k+1} - x_*|| \leq C_1\ \delta\ ||x_k - x_*||. \tag{4.28}$$

According to (4.24), (4.23) and (4.20), y_{k-1} belongs to ω and according to (4.25), (4.23) and (4.20), x_k belongs to ω. Therefore y_k is well defined by (4.8) and we have (4.19). This inequality and (4.25) show (4.27). Now, according to (4.27), (4.23) and (4.20), y_k belongs to ω. In the same way, $\bar{y}_k$ belongs to ω. So, $\bar{x}_{k+1}$ and x_{k+1} are well defined by (4.6) and (4.9) respectively. The inequalities (4.26), (4.22) and (4.21) show that G_k is nonsingular and satisfies

$$||G_k^{-1}|| \leq \frac{1}{\frac{1}{\eta} - \bar{\delta}}, \tag{4.29}$$

from which we deduce with (4.21), (4.26) and $G_*^{-1}-G_*^{-1} = G_k^{-1}\ (G_*-G_k)\ G_*^{-1}$:

$$||G_k^{-1} - G_*^{-1}|| \leq ||G_k^{-1}||\ ||G_*^{-1}||\ ||G_k - G_*|| \leq C_2\ \delta.$$

Let C_3 be the constant given by lemma 4.1. Then, Taylor developments give easily the following inequalities :

$$||x_{k+1} - x_*|| \le ||x_{k+1} - \bar{x}_{k+1}|| + ||\bar{x}_{k+1} - x_*||,$$

$$||x_{k+1} - \bar{x}_{k+1}|| \le C_4 ||y_k - \bar{y}_k|| + C_5 ||y_k - x_*||^2 + C_2 C_6 \delta ||y_k - x_*||,$$

$$||y_k - \bar{y}_k|| \le C_7 ||y_{k-1} - x_*|| \, ||x_k - x_*||.$$

By combining those inequalities with (4.7), (4.24), (4.25) and (4.19), we get (4.28) with $C_1 = C_3 + C_4C_7C_0 + C_5C_0^2 + C_2C_6C_0$. By definition, x_k is obtained from y_{k-1} by (4.9) and the development of $g(y_{k-1})$ around x_* gives with (4.18) and (4.29) :

$$(4.30) \qquad ||x_k - x_*|| \le (1 + C_8) \, ||y_{k-1} - x_*||.$$

Now, we can prove the theorem with $C := \min(\bar{\delta}, 1/C_1)/(1+C_8)$. Indeed, if (4.15) and (4.16) are satisfied, we see by (4.30) that (4.24)-(4.26) are verified for $k = 1$ with $\delta = (1+C_8)Cr \le \bar{\delta}$. Then, $C_1\delta \le r$ and (4.28) shows that (4.17) is satisfied for $k = 1$. The fact that r is less than 1 and (4.27) allow to use a recurrence argument. So, (4.17) is also satisfied for k greater than 1 •

We shall say that two positive real sequences (α_k) and (β_k) are <u>equivalent</u> (we note $\alpha_k \sim \beta_k$) if $\alpha_k = O(\beta_k)$ and $\beta_k = O(\alpha_k)$. The next proposition gives two useful equivalences.

<u>Proposition 4.3</u> : <u>Suppose that assumptions</u> A <u>and</u> B <u>are satisfied</u>. <u>Let</u> (G_k) <u>be a bounded sequence of nonsingular matrices of order</u> n-m <u>such that</u> (G_k^{-1}) <u>be also bounded</u>. <u>Let</u> (x_k) <u>in</u> ω <u>and</u> (y_k) <u>in</u> ω <u>be the sequences generated by the algorithm</u> (4.8)-(4.9) <u>starting from a point</u> y_0 <u>in</u> ω. <u>If</u> (y_k) <u>converges to a solution</u> x_* <u>of</u> (1.1), <u>we have</u>

$$(4.31) \qquad ||d_k|| \sim ||x_k - x_*||,$$

$$(4.32) \qquad ||e_k|| \sim ||y_k - x_*||.$$

<u>Proof</u>. If (y_k) converges to x_* then so does (x_k). From the definition (4.10) of r_k and the development of $c(x_k)$ around x_*, we get

(4.33) $$r_k = - A_*^- A_* (x_k - x_*) + o(||x_k - x_*||).$$

Then, by using the identity (3.3), we obtain

(4.34) $$y_k - x_* = Z_*^- Z_* (x_k - x_*) + o(||x_k - x_*||).$$

By using the boundedness of (G_k^{-1}) and (4.2), we see that $t_k = -Z_*^- G_k^{-1} Z_*^{-T} L_* (y_k - x_*) + o(||y_k - x_*||)$. Eventually,

$$d_k = - [A_*^- A_* + Z_*^- G_k^{-1} G_* Z_*] (x_k - x_*) + o(||x_k - x_*||).$$

This estimate shows that $d_k = O(||x_k - x_*||)$. To prove the converse, we only have to show that the operator in square brackets is nonsingular with bounded inverse. If it were not the case, there would exist a subsequence K of subscripts and a sequence $(\xi_k : k \in K)$ in $\mathbb{R}^n$ such that

(4.35) $$||\xi_k|| = 1 \quad \text{for } k \text{ in } K,$$

(4.36) $$[A_*^- A_* + Z_*^- G_k^{-1} G_* Z_*] \xi_k \longrightarrow 0 \quad \text{for } k \text{ in } K.$$

By multiplying (4.36) by A_* (resp. Z_*), we should obtain $A_* \xi_k \longrightarrow 0$ (resp. $G_k^{-1} G_* Z_* \xi_k \longrightarrow 0$). We should deduce $Z_* \xi_k \longrightarrow 0$ because of the boundedness of (G_k) and the nonsingularity of G_*. Finally, with (3.3), we should have $\xi_k \longrightarrow 0$ that would contradict (4.35). So, (4.31) is proved. The proof of (4.32) is similar and is based on the estimate

$$e_k = - [A_*^- A_* + Z_*^- G_k^{-1} Z_*^{-T} L_*] (y_k - x_*) + o(||y_k - x_*||) \bullet$$

5 - Conditions for q-superlinear convergence

The theorem 4.2 has an immediate corollary which states that if in addition to (4.16), the sequence (G_k) converges to G_* then (x_k) converges to x_* q-superlinearly (see for example the argument in the proof of corollary 3.5 in Han (1976)). However, this assumption on (G_k) is usually not satisfied in practice, in particular when those matrices are generated by quasi-Newton formula. Assuming that (x_k) converges to x_*, the next theorem gives necessary and sufficient conditions on (G_k) in order to have the q-superlinear convergence of

(x_k). It is the analogue of the theorem 2.2 of Denis-Moré (1974) valid for quasi-Newton methods without constraints.

Theorem 5.1 : Suppose that assumptions A and B are satisfied and that (y_k) and (x_k) are generated in ω from a point y_0 by the algorithm (4.8)-(4.9) with a sequence (G_k) of nonsingular matrices. Suppose that (y_k) converges to x_* (then so does (x_k)). Then, the following statements are equivalent :

(i) (x_k) converges q-superlinearly,
(ii) $g(y_{k+1}) = o(||x_k - x_*||)$,
(iii) $(G_k - G_*)\, Z(y_k)\, t_k = o(||x_k - x_*||)$.

Proof. The estimate (4.34) shows that $y_k - x_* = O(||x_k - x_*||)$. It is then easy to get (we use $A_* Z_*^- = 0$)

$$A_* (x_{k+1} - x_*) = o(||x_k - x_*||). \tag{5.1}$$

According to (3.3), it remains to estimate $Z_*(x_{k+1} - x_*)$. This will depend on the quality of the tangent step t_k. Let us first prove the equivalence (i) <=> (ii). With (4.2) and (4.34), we have

$$\begin{aligned} g(y_{k+1}) &= Z_*^{-T} L_* (y_{k+1} - x_*) + o(||y_{k+1} - x_*||) \\ &= G_* Z_* (x_{k+1} - x_*) + o(||x_{k+1} - x_*||). \end{aligned}$$

Then (ii) is clear from (i). If (ii) is satisfied, this estimate and the nonsingularity of G_* give

$$Z_* (x_{k+1} - x_*) = o(||x_{k+1} - x_*||) + o(||x_k - x_*||).$$

This estimate, (5.1) and the identity (3.3), shows (i). Now, let us show that in any of the situations (i), (ii) or (iii), we have

$$t_k = O(||x_k - x_*||). \tag{5.2}$$

This estimate is clear when (G_k^{-1}) is bounded, but we do not suppose this here. By writing $t_k = (x_{k+1} - x_*) - (y_k - x_*)$ and using (4.34), we see that (5.2) is clearly satisfied when (i) is true and therefore when (ii) is true. When (iii) is satisfied, we have

$$G_* Z(y_k) t_k = G_k Z(y_k) t_k + o(||x_k-x_*||) = -g(y_k) + o(||x_k-x_*||).$$

Then by developing $g(y_k)$ around x_* and by using (4.34) and the nonsingularity of G_*, we get

$$Z(y_k) t_k = O(||x_k-x_*||).$$

But $t_k = Z(y_k)^- Z(y_k) t_k$, so (5.2) is satisfied. Now, from (5.2), it follows that $x_{k+1}-x_* = O(||x_k-x_*||)$, $y_{k+1}-y_k = O(||x_k-x_*||)$ and, with (4.33) and (5.1),

$$r_{k+1} = o(||x_k-x_*||). \tag{5.3}$$

It remains to show the equivalence (ii) <=> (iii). By developing $g(y_{k+1})$ around y_k and by using (4.2) and (5.3), we have

$$g(y_{k+1}) = g(y_k) + Z_*^{-T} L_* t_k + o(||x_k-x_*||).$$

Because $g(y_k) = - G_k Z(y_k) t_k$ and $t_k = Z(y_k)^- Z(y_k) t_k = Z_*^- Z(y_k) t_k + o(||x_k-x_*||)$, we obtain

$$g(y_{k+1}) = - (G_k-G_*) Z(y_k) t_k + o(||x_k-x_*||).$$

The equivalence (ii) <=> (iii) follows •

In the statement (ii) of theorem 5.1, $g(y_{k+1})$ can be replaced by $g(x_{k+1})$, but the reduced gradient is not calculated at x_{k+1} in the algorithm. The statement (iii) is equivalent to

$$(G_k^{-1}-G_*^{-1}) g(y_k) = o(||x_k-x_*||)$$

which is based on the gap between the inverse of the hessians. The statement (iii) can also be replaced by many other equivalent estimates. For example,

$$(G_k-G_*) Z_* (x_k-x_*) = o(||x_k-x_*||).$$

The advantage of (iii) is that it does not require the boundedness of the sequences (G_k) or (G_k^{-1}). If this is assumed, proposition 4.3 shows that the estimates can be done in relation to $||d_k||$ rather than $||x_k-x_*||$.

The conditions (4.16) and (iii) of theorem 5.1 show the advantage of the reduced quasi-Newton methods over the SQP method with regard to the approximation of the hessian of the lagrangian. Indeed, a necessary and sufficient condition for the SQP method to generate q-superlinearly convergent sequences is that

$$Z_*^{-T} (L_k - L_*) (x_k - x_*) = o(||x_k - x_*||),$$

where L_k is the updated approximation of L_*. Therefore, in the SQP method, the (n-m) x n matrix $Z_*^{-T} L_*$ has to be correctly approximated and not only $Z_*^{-T} L_* Z_*^-$ like in the reduced methods. This famous result can be found in Boggs-Tolle-Wang (1982) and Nocedal-Overton (1985).

The statement (iii) of theorem 5.1 shows that the superlinear convergence depends on the quality of the approximation of G_* by G_k in certain directions. In fact, a stronger condition than (iii) will be satisfied when concrete update schemes are considered : see Gilbert (1986,a) for a detailed analysis.

6 - Globalization of the algorithm

In order to globalize the local algorithm (4.8)-(4.9), we introduce a step-size parameter ρ_k. For that, we consider the following exact penalty function :

$$\Theta_p(x) = f(x) + p\, ||c(x)||_1, \tag{6.1}$$

where p is the positive penalty parameter and $||.||_1$ is the 1-norm on $\mathbb{R}^m$. If p is taken greater than $||\lambda_*||_\infty$ (where $||.||_\infty$ is the sup-norm on $\mathbb{R}^m$), feasible minimizers of (1.1) and (6.1) are the same (see Fletcher (1981) for example). Other norms than the 1-norm can be used in (6.1) : see Bonnans-Gabay (1984). In order to minimize Θ_p, we need to calculate descent directions of this non-differentiable function. On that point, a crucial observation has been made by Han (1977) : the displacement d_k^{SQP} of the SQP method is a descent direction of Θ_p at x_k (if some natural hypothesis are satisfied). Therefore a better approximation x_{k+1} of the solution x_* will be obtained by taking

$$x_{k+1} = x_k + \rho_k d_k^{SQP},$$

where ρ_k gives the step-size and is obtained from some rule using Θ_p as a "merit" function. Let us try to use the same globalizing technique for our algorithm. Is there any descent direction of Θ_p among the displacements r_k, t_k, d_k and e_k given by (4.10)-(4.13) ? The inconvenient of r_k and therefore of d_k and e_k is that this displacement is calculated by using two different points y_{k-1} and x_k that can be far from each other when x_k is far from x_*. So, it is difficult to see when those directions are descent directions for Θ_p. On the other hand, t_k uses only the point y_k in its definition and if G_k is positive definite, it is certainly a descent direction of Θ_p at y_k. Indeed, the displacement is tangent to $c^{-1}(c(y_k))$ at y_k and $f'(y_k).t_k$ is negative. Therefore at the first order, the first term of the right hand side of (6.1) will decrease rather than the second term will remain constant. Those remarks lead us to define a descent arc of Θ_p at y_k, tangent to t_k :

$$y_k(\rho) = y_k + \rho\, t_k + \rho^a\, r_{k+1} \quad , \quad a > 1. \tag{6.2}$$

Search arcs have already been proposed by Maine-Polak (1982) to cope with the Maratos effect of the SQP method (see further) and by Gabay (1982,b), also to avoid the Maratos effect for his algorithm. This globalizing technique based on the arc (6.2) gives the priority to the minimizing step t_k and this is due to the asymmetry of the local method (4.8)-(4.9). This priority can be harmful in some circumstances but it can be suppressed by adding a restoration step to the local method (see Gilbert (1986,b)). The point y_{k+1} is then obtained from y_k by a particular choice of ρ :

$$y_{k+1} := y_k(\rho_k). \tag{6.3}$$

The step-size ρ_k is determined so that the following Armijo-like criterion is satisfied

$$\beta \in]0,1[, \tag{6.4}$$

$$\rho_k := \beta^{l_k}, \tag{6.5}$$

where l_k is the smallest non-negative integer such that

$$\theta_p(y_k(\beta^{l_k})) \le \theta_p(y_k) + \beta^{l_k}\,\alpha\, f'(y_k).t_k$$

(6.6)

$$- \beta^{al_k}\,\alpha\,(p - ||\lambda(y_k)||_\infty)\;||c(y_k)||_1.$$

In this inequality, α is a real number chosen in $]0,1/2[$ for reasons that will be clear at the end of this section. The exponent (al_k) of β in the last term of (6.6) takes into account the curvature of the search path (6.2). The vector $\lambda(y_k)$ is the approximation at y_k of the Lagrange multiplier λ_* and is defined by

(6.7) $$\lambda(y) := - A(y)^{-T}\,\nabla f(y).$$

It is just the first term of (4.14). So, usually $(\lambda(y_k))$ will not converge superlinearly. Unfortunately, in the last term of (6.6), we have to use the value of the constraints at the point y_k and not at the point

(6.8) $$x_{k+1} := y_k + t_k$$

which is used to calculate r_{k+1}. This seems necessary to have $\rho_k = 1$ assymptotically (see theorem 6.3). The Armijo rule (6.4)-(6.7) has the advantage to give to ρ_k the value 1 as soon as it is accepted by the criterion (6.6) and so not to prevent the q-superlinear convergence of the sequence (x_k). Now, we have to examine in what conditions the inequality (6.6) can be realized by taking l_k great enough. This is the subject of the following lemma.

Lemma 6.1 : *Suppose that assumptions A and B are satisfied and that a point y_k is given in ω such that x_{k+1} and $x_{k+1} + r_{k+1}$ will also be in ω. Suppose that α is in $]0,1[$ and that there exists positive constants $\underline{p}$, $\bar{p}$, $\underline{h}$ and $\bar{h}$ such that*

(6.9)
$$\underline{p} + ||\lambda(y_k)||_\infty \le p \le \bar{p},$$
$$(G_k^{-1} - \underline{h}\, I) \text{ will be symmetric positive definite},$$
$$||G_k^{-1}|| \le \bar{h}.$$

Then the rule (6.4)-(6.6) allows to determine a positive step-size ρ_k. Moreover, if M is a positive constant such that

(6.10) $$||c(y_k)||_1 \le M \quad \text{and} \quad \rho_k\,||g(y_k)||^2 \le M,$$

then, there exists a positive real $\underline{\rho}$ that depends only on f, c, $\underline{p}$, $\bar{p}$, $\underline{h}$, $\bar{h}$, α, β and M such that

$$\rho_k \geq \underline{\rho} > 0.$$

Proof. By using

$$c^i(x_{k+1}) = c^i(y_k) + \frac{1}{2}\,(c^i)''(u_{i,k}).(t_k)^2, \quad i = 1, \ldots, m, \tag{6.11}$$

where the points $u_{i,k}$ are in the convex set ω, the developments of f and c^i around y_k write

$$\begin{aligned} f(y_k+\rho t_k+\rho^a r_{k+1}) &= f(y_k) + \rho f'(y_k).t_k + \rho^a(\lambda(y_k),c(y_k)) \\ &\quad + \frac{\rho^a}{2} \sum_{i=1}^{m} \lambda^i(y_k)(c^i)''(u_{i,k}).(t_k)^2 + \frac{1}{2}\, f''(z_k(\rho)).(\rho t_k+\rho^a r_{k+1})^2, \end{aligned} \tag{6.12}$$

$$\begin{aligned} c^i(y_k+\rho t_k+\rho^a r_{k+1}) &= (1-\rho^a)c^i(y_k) - \frac{\rho^a}{2}\,(c^i)''(u_{i,k}).(t_k)^2 \\ &\quad + \frac{1}{2}\,(c^i)''(v_{i,k}(\rho)).(\rho t_k+\rho^a r_{k+1})^2, \end{aligned} \tag{6.13}$$

where the points z_k and $v_{i,k}$ depend on ρ and are in ω. From those developments and by supposing ρ in $]0,1]$, we get

$$\begin{aligned} \Theta_p(y_k+\rho t_k+\rho^a r_{k+1}) &\leq \Theta_p(y_k) + \rho f'(y_k).t_k \\ &\quad - \rho^a(p-||\lambda(y_k)||_\infty)||c(y_k)||_1 \\ &\quad + C_1(\rho^a+\rho^2)||t_k||^2 + C_2\rho^{2a}||r_{k+1}||^2, \end{aligned} \tag{6.14}$$

where C_1 and C_2 are positive constants that depend only on f,c and $\bar{p}$. From (6.11), we have

$$||r_{k+1}|| \leq C_3\,||c(y_k)||_1 + C_4\,||t_k||^2, \tag{6.15}$$

where C_3 and C_4 depend only on c. Now, suppose that (6.6) is not true for a given ρ in $]0,1]$. Then, by using (6.14), $-f'(y_k).t_k = (G_k^{-1}g(y_k),g(y_k)) \geq \underline{h}||g(y_k)||^2$ and (6.15), we obtain

$$\begin{aligned} &\rho\,||g(y_k)||^2 + \rho^a\,||c(y_k)||_1 \\ &\quad \leq C_5\,(\rho^a+\rho^2)\,||g(y_k)||^2 + C_6\,\rho^{2a}\,||c(y_k)||_1^2 + C_7\,\rho^{2a}\,||g(y_k)||^4, \end{aligned} \tag{6.16}$$

where C_5, C_6 and C_7 are positive constants that depend only on f, c, $\underline{p}$, $\bar{p}$, α, $\underline{h}$ and $\bar{h}$. This inequality shows that ρ cannot be arbitrarily small if $||g(y_k)||+||c(y_k)||_1 \neq 0$. This proves the first part of the lemma. If the rule (6.4)-(6.6) gives a step-size ρ_k smaller than 1, (6.6) is not satisfied with $\rho = \rho_k/\beta$ and the inequality (6.16) with (6.10) gives :

$$\rho_k \, ||g(y_k)||^2 + \rho_k^a \, ||c(y_k)||_1 \leq C_8 \, \rho_k^b \, [\rho_k \, ||g(y_k)||^2 + \rho_k^a \, ||c(y_k)||_1],$$

where b := min(1,a-1) and C_8 depends only on f, c, $\underline{p}$, $\bar{p}$, α, β, $\underline{h}$, $\bar{h}$ and M. Because $\rho_k \, ||g(y_k)||^2 + \rho_k^a \, ||c(x_{k+1})||_1 \neq 0$ (otherwise $\rho_k = 1$), the last inequality proves the second part of the lemma •

The inequality (6.9) shows that the penalty parameter has to be large enough to ensure the decrease of θ_p along the arc (6.2) and that its lower bound depends on the current point y_k. So, sometimes it will be necessary to update the penalty parameter that we shall note p_k. We shall suppose that the adapting rule of p_k will satisfy the following three conditions :

(6.17) $\quad p_k \geq ||\lambda(y_k)||_\infty + \underline{p}$, for every k,

(6.18) there exists a subscript K such that for every k greater than K, ($p_{k-1} \geq ||\lambda(y_k)||_\infty + \underline{p}$) implies that $p_k = p_{k-1}$,

(6.19) (p_k) is bounded if and only if p_k is modified finitely often.

In (6.17) and (6.18), $\underline{p}$ is a given positive constant. The condition (6.18) means that eventually (for $k \geq K$), p_k is modified only if it is necessary to have (6.17). So $(p_k : k \geq K)$ is an increasing sequence. An example of adapting rule satisfying those conditions is given in Mayne-Polak (1982) :

<u>if</u> $p_{k-1} \geq ||\lambda(y_k)||_\infty + \underline{p}$

<u>then</u> $p_k := p_{k-1}$

<u>else</u> $p_k := \max (\delta \, p_{k-1} \, , \, ||\lambda(y_k)||_\infty + \underline{p})$,

where δ is a given constant greater than 1 in order to satisfy (6.19). We are now able to state the algorithm which globalizes the local method (4.8)-(4.9).

Algorithm RQN :

1 - Choose a convergence tolerance ε, β in $]0,1[$, α in $]0,1/2[$ and $a > 1$.

2 - Choose y_0 in ω and a positive definite symmetric matrix G_0 of order n-m.

3 - Let $k := 0$.

4 - Repeat :

4.1 - Linearize the constraints at y_k : choose a right inverse $A(y_k)^-$ of $\nabla c(y_k)$ and a basis $Z(y_k)^-$ of $N(\nabla c(y_k))$.

4.2 - Evaluate $\lambda(y_k) := - A(y_k)^{-T}\nabla f(y_k)$ and $g(y_k) := Z(y_k)^{-T}\nabla f(y_k)$.

4.3 - If $k \geq 1$ then evaluate the symmetric matrix G_k by updating G_{k-1}.

4.4 - Tangent step : evaluate $t_k := -Z(y_k)^- G_k^{-1} g(y_k)$ and $x_{k+1} := y_k + t_k$.

4.5 - Restoration step : evaluate $c(x_{k+1})$ and $r_{k+1} := -A(y_k)^- c(x_{k+1})$.

4.6 - If $||g(y_k)|| + ||c(x_{k+1})|| < \varepsilon$ then stop.

4.7 - Adapt p_k according to (6.17)-(6.19).

4.8 - Search a point y_{k+1} from y_k along the arc (6.2) in order to decrease the penalty function (6.1) (with $p = p_k$) according to (6.3)-(6.6).

In the partitioned framework (see section 3), only one linear system has to be solved at the step 4.2. Indeed, if $A(y_k) = [C(y_k)\ D(y_k)]$, $\lambda(y_k)$ is obtained by solving

$$C(y_k)^T \lambda(y_k) = - \nabla f(y_k)^{(1)},$$

where $\nabla f(y_k)^{(1)}$ are the first m components of $\nabla f(y_k)$. Then $g(y_k) = D(y_k)^T \lambda(y_k) + \nabla f(y_k)^{(2)}$ where $\nabla f(y_k)^{(2)}$ are the last n-m components of $\nabla f(y_k)$. This is known as the adjoint state method to compute the reduced gradient. A crucial point of the algorithm that has only been mentioned at the step 4.3 concerns the update of the matrices G_k. This update scheme is expected to generate a bounded sequence (G_k^{-1}) of uniformly positive definite symmetric matrices G_k^{-1}. That is to say that there exists a positive constant C such that for every v in $\mathbb{R}^{n-m}$ and every subscript k, we have :

$$(G_k^{-1} v, v) \geq C\ ||v||^2.$$

This property is really not easy to show. By using the same type of arguments that are used in unconstrained optimization, we were only able to prove it in a local framework (when (x_0, G_0) is supposed to be close to (x_*, G_*)) or when a little bit less than the q-linear convergence of (x_k) is assumed (see Gilbert (1986,a)). The next theorem analyses the global convergence of the algorithm RQN under that hypothesis.

Theorem 6.2 : Suppose that assumptions A and B are satisfied and that f is bounded from below on ω. Let (x_k), (y_k) and (G_k), the sequences generated by the algorithm RQN with α in $]0,1[$. Suppose that (x_k) and (y_k) are in ω and that (G_k^{-1}) is bounded and uniformly positive definite. Then, either (p_k) is unbounded and $(y_k : p_k \neq p_{k-1})$ has no accumulation point, or (p_k) is bounded and

$$(6.20) \qquad ||g(y_k)|| + ||c(y_k)||_1 \longrightarrow 0.$$

Proof. Suppose first that (p_k) is unbounded and let K be the subsequence of those subscripts $k \geq K$ for which $p_k \neq p_{k-1}$. By (6.19), K is unbounded and by (6.18) :

$$p_{k-1} < ||\lambda(y_k)||_\infty + \underline{p}$$

for k in K. Because $(p_k : k \geq K)$ is an increasing sequence, we see that $||\lambda(y_k)||_\infty \longrightarrow \infty$ for $k \longrightarrow \infty$ in K. Therefore $(y_k : p_k \neq p_{k-1})$ has no accumulation point (here, we use the continuity of $y \longrightarrow \lambda(y)$ and so, the surjectivity of $\nabla c(y)$ and assumption B are strongly invoked). Now, let us suppose that (p_k) is bounded. From (6.19), p_k is constant for k great enough. Let us say that $p_k = p$ for $k \geq K_1$. So, at each iteration the same penalty function Θ_p decreases. The function f being bounded from below, we get

$$p\,||c(y_k)||_1 \leq \Theta_p(y_{K_1}) - \inf f \ , \quad \text{for } k \geq K_1 .$$

Therefore, $(||c(y_k)||_1)$ is bounded. On the other hand, from (6.6), (6.17) and the existence of a positive constant $\underline{h}$ such that $(G_k^{-1} - \underline{h}I)$ is positive definite, we find :

$$\rho_k \propto \underline{h}\,||g(y_k)||^2 + \rho_k^a \propto \underline{p}\,||c(y_k)||_1 \leq \Theta_p(y_k) - \Theta_p(y_{k+1}) \ , \quad \text{for } k \geq K_1 . \tag{6.21}$$

Because $(\Theta_p(y_k))$ converges (a decreasing bounded from below sequence), this inequality certainly shows that $(\rho_k||g(y_k)||^2)$ is bounded. Then, we can apply lemma 6.1 that states the existence of a positive lower bound for (ρ_k). By taking the limit in k in (6.21), we obtain the result •

A last problem to tackle concernes the question of the admissibility of the unity step-size. When $\rho_k = 1$ is accepted by (6.6), the algorithm RQN proceeds like the local method (4.8)-(4.9) and q-superlinear convergence of (x_k) can occur if the reduced hessian G_* is correctly approched by G_k (see theorem 5.1, statement (iii)). It is known that this admissibility property is not satisfied when the SQP method is globalized with the penalty function (6.1) and the technique described at the beginning of this section. This has been called the "Maratos effect" of the SQP method (see Maratos (1978)) and several remedies have been proposed to overcome that drawback (see Chamberlain-Lemarechal-Pedersen-Powell (1982), Mayne-Polak (1982) and Bonnans (1984)). That inconvenience is not shared with our algorithm. In fact, when $c(y_k)=0$ (a favorable situation for the appearance of the Maratos effect), the total displacement $e_k = t_k + r_{k+1}$ is exactly the same as the one of the SQP method with the Mayne-Polak's correction. This may explain that.

Let (x_k) in ω, (y_k) in ω and (G_k) be the sequences generated by the algorithm RQN and suppose that (y_k) converges to a solution x_* of (1.1). We are interested in finding conditions under which ρ_k will be equal to 1 for all but finitely many subscripts k in the subsequence considered. Let K be a subsequence of subscripts. The following four properties will be meaningful :

(6.22) $\quad t_k = O(||r_{k+1}||)$ for k in K,

(6.23) $\quad (G_k - G_*) Z_* t_k = o(||t_k||)$ for k in K,

(6.24) $\quad ||G_k - G_*|| \leq M$ for k in K,

(6.25) $\quad \rho_k < 1$ and $t_k = o(||r_k||)$ for k in K.

The property (6.23) concerns the approximation of the reduced hessian G_* by G_k and recalls the condition (iii) of theorem 5.1 that writes when (G_k) and (G_k^{-1}) are bounded :

(6.26) $\quad (G_k - G_*) Z_* t_k = o(||d_k||)$.

Therefore, (6.23) is usually stronger than (6.26) and, in fact, is satisfied by some subsequences of subscripts when (G_k) is updated by the BFGS formula (see Gilbert (1986,a)). The property (6.24) is very strong when M is small and is usually not satisfied when second order derivatives are not calculated.

The next theorem shows that for the subsequences K for which (6.22) or (6.23) or (6.24) with M small enough is satisfied, the rule (6.3)-(6.6) will give $\rho_k = 1$ for all but finitely many k in K. When the property (6.25) is satisfied the same conclusion does not hold any more in the framework of the global algorithm RQN. Nevertheless, the result obtained below with (6.25) has allowed to show that in the concrete algorithms in Gilbert (1986,a), the subsequences satisfying (6.25) also admit a unity step-size after a finite number of iterations.

Theorem 6.3 : Suppose that assumptions A and B are satisfied. Let (x_k), (y_k) and (G_k) be the sequences generated by the algorithm RQN with α in $]0,1/2[$. Suppose that (x_k) and (y_k) are in ω, that $(||G_k^{-1}||)$ is bounded by $\bar{h}$, that there exists a positive constant $\underline{h}$ such that $(G_k^{-1} - \underline{h}I)$ is positive definite and that (y_k) converges to x_*. Let K be a subsequence of subscripts. Then,

(i) if (6.22) or (6.23) is satisfied then $\rho_k = 1$ for all but finitely many k in K,

(ii) there exists a positive constant $\bar{M}$ that depends only on c, α, $\underline{h}$ and $\bar{h}$ such that if (6.24) is satisfied with $M \leq \bar{M}$ then $\rho_k = 1$ for all but finitely many k in K,

(iii) if (6.25) is satisfied, then $r_{k+1} = o(||r_k||\ ||t_k||)$ for k in K.

Proof. Since (y_k) converges to x_*, (6.18) and (6.19) imply that p_k is modified finitely often. So, we shall suppose that $p_k = p$ for all k. We develop $\theta_p(y_k+t_k+r_{k+1})$ around y_k at the second order in t_k and the first order in r_{k+1}. The development (6.12) gives

$$f(y_k+e_k) = f(y_k) + f'(y_k).t_k + (\lambda(y_k),c(y_k)) + \frac{1}{2}(L_*t_k,t_k) + o(||t_k||^2) + o(||r_{k+1}||),$$

while (6.13) shows that $c(y_k+e_k) = o(||t_k||^2) + o(||r_{k+1}||)$. If we note

$$\Delta_k := f'(y_k).t_k - (p - ||\lambda(y_k)||_\infty)\ ||c(y_k)||_1$$

which is negative, we obtain

$$\theta_p(y_k+e_k) \leq \theta_p(y_k) + \Delta_k + \frac{1}{2}(L_*t_k,t_k) + o(||t_k||^2) + o(||r_{k+1}||).$$

But $t_k = \bar{Z}_*Z_*t_k + o(||t_k||)$ and the boundedness of (G_k) allows to write $g(y_k) = -G_kZ_*t_k + o(||t_k||)$. Therefore, by using $f'(y_k).t_k = -(G_k^{-1}g(y_k),g(y_k)) = -(G_kZ_*t_k,Z_*t_k) + o(||t_k||^2)$, we obtain

$$\theta_p(y_k+e_k) - \theta_p(y_k) - \alpha\,\Delta_k \leq (\tfrac{1}{2} - \alpha)\,\Delta_k - \frac{1}{2}((G_k-G_*)Z_*t_k,Z_*t_k) + o(||t_k||^2) + o(||r_{k+1}||). \tag{6.27}$$

We have to find in what conditions the left hand side of (6.27) is non-positive. Suppose it is positive when k belongs to the subsequence K. Then, by using the inequality $c_1||t_k|| \leq ||g(y_k)||$ (c_1 is a positive constant that depends only on c and $\bar{h}$), the uniform positive definiteness of (G_k^{-1}), (6.17) and (6.11), we obtain :

$$\underline{h}\,c_1^2\,||t_k||^2 + p\,||c(y_k)||_1 \leq \frac{-1}{1-2\alpha}((G_k-G_*)Z_*t_k,Z_*t_k) + o(||t_k||^2) + o(||c(y_k)||_1).$$

This leads to a contradiction if (6.22) or (6.23) or (6.24) with $M < (1-2\alpha)\ \underline{h}\ c_1^2/||Z_*||^2$ is verified. When (6.25) is satisfied, the left hand side of (6.27) is positive and by using the last inequality, we obtain :

$$||t_k||^2 + ||c(y_k)||_1 = o(||r_k||\ ||t_k||).$$

Then, the definition of r_{k+1} and (6.11) give the result •

7 - Conclusion

We have studied in this paper the local and global convergence of a variable metric algorithm for equality constrained optimization in which the order of the updated matrices is n-m. This reduced method can be seen as making a link between GRG-like methods which are feasible methods ($c(x_k) = 0$ for all k) with reduced metrics (of order n-m) and the SQP method which is an unfeasible method with full metrics (of order n) : the studied algorithm is indeed an unfeasible method with reduced metrics. The algorithm inherits also the good properties of both methods (reduced metrics, superlinear convergence and unfeasibility) and shows, in particular, that locally only one restoration step is necessary to obtain the superlinear convergence in GRG-like methods if the reduced metrics is correctly approximated.

The global convergence is obtained by Han's technique to globalize the SQP method. The l_1 penalty function is used as merit function and is decreased along arc-shaped search path.

An important facet of the method has not been tackled here and is reported somewhere else (Gilbert (1986,a)). This concerns the update of the reduced matrices G_k. This one is based on a secant equation using the reduced gradient g. The fact that the gradient of g at x_* (see (4.2)) is not equal to G_* (and cannot be because $\nabla g(x_*)$ is an (n-m) x n matrix while G_* is of order n-m) leads to an alternative. Either the reduced gradient is evaluated twice per iteration, at y_k and x_{k+1}, or it is evaluated only once per iteration, at y_k. In the first case, $g(y_k)$ and $g(x_{k+1})$ are used in the secant equation and the matrices G_k are updated at each iteration but with the inconvenient to have to

linearize the constraints twice per iteration : see Coleman-Conn (1984) and Gilbert (1986,a). In the second case, $g(y_k)$ and $g(y_{k+1})$ are used in the secant equation but the matrices G_k are not usually updated at each iteration. An update criterion has to be introduced in order to decide when an update is appropriate. Despite of this, the superlinear convergence can be achieve either in a local framework (see Nocedal-Overton (1985) for the algorithm (1.12)-(1.14)) or in a global framework (see Gilbert (1986,a) for the algorithm RQN of section 6).

Acknowledgments

This work was supported in part during 1984-1985 at the "Centre d'Etudes Nucléaires", Fontenay-aux-Roses (France) under a grant from the Commission of the European Community and in part at the "Institut National de Recherche en Informatique et en Automatique" (INRIA), Le Chesnay (France). I would like to thank those organisms and to express my gratitude to J.F. Bonnans for helpful discussions and suggestions and for his friendly and constant advise.

References

J. Abadie, J. Carpentier (1969). Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints. Optimization, R. Fletcher, ed., Academic Press, London.

A. Ben-Israel, T.N.E. Greville (1974). Generalized inverses : theory and applications. John Wiley & Sons.

J. Blum, J.Ch. Gilbert, B. Thooris (1985). Parametric identification of the plasma current density from the magnetic measurements and the pressure profile, code IDENTC. Report of JET contract number JT3/9008. Centre d'Etudes Nucléaires (DRFC), B.P. 6, 92265 Fontenay-aux-Roses (France).

P.T. Boggs, J.W. Tolle, P. Wang (1982). On the local convergence of quasi-Newton methods for constrained optimization. SIAM Journal on Control and Optimization 20/2, 161-171.

J.F. Bonnans (1984). Asymptotic admissibility of the unity stepsize in exact penalty methods I: equality-constrained problems. Rapport de recherche de l'INRIA RR-273. 78153 Le Chesnay, France.

J.F. Bonnans, D. Gabay (1984). Une extension de la programmation quadratique successive. Lecture Notes in Control and Information Sciences 63, 16-31. A. Bensoussan, J.L. Lions, (eds.). Springer-Verlag.

R.H. Byrd (1985). An example of irregular convergence in some constrained optimization methods that use the projected hessian. Mathematical Programming 32, 232-237.

R.H. Byrd, R.B. Schnabel (1986). Continuity of the null space basis and constrained optimization. Mathematical Programming 35, 32-41.

R.M. Chamberlain, C. Lemaréchal, H.C. Pedersen, M.J.D. Powell (1982). The watchdog technique for forcing convergence in algorithms for constrained optimization. Mathematical Programming Study 16, 1-17.

T.F. Coleman, A.R. Conn (1982,a). Nonlinear programming via an exact penalty function: asymptotic analysis. Mathematical Programming 24, 123-136.

T.F. Coleman, A.R. Conn (1982,b). Nonlinear programming via an exact penalty function: global analysis. Mathematical Programming 24, 137-161.

T.F. Coleman, A.R. Conn (1984). On the local convergence of a quasi-Newton method for the nonlinear programming problem. SIAM Journal on Numerical Analysis 21/4, 755-769.

Y.M. Danilin, B.N. Pschenichny (1965). Numerical methods in extremal problems. MIR, Moscow. (English translation, 1978).

J.E. Dennis, J.J. Moré (1977). Quasi-Newton methods, motivation and theory. SIAM Review 19, 46-89.

R. Fletcher (1977). Practical methods of optimization. Vol. 2 : Constrained optimization. Wiley.

D. Gabay (1975). Efficient convergence of implementable gradient algorithms and stepsize selection procedures for constrained minimization. International Computing Symposium 1975. F. Gelenbe, K.D. Potier, eds.. North-Holland Publishing Company.

D. Gabay (1982,a). Minimizing a differentiable function over a differentiable manifold. Journal of Optimization Theory and its Application 37/2, 171-219.

D. Gabay (1982,b). Reduced quasi-Newton methods with feasibility improvement for nonlinearly constrained optimization. Mathematical Programming Study 16, 18-44.

D. Gabay, D.G. Luenberger (1976). Efficiently converging minimization methods based on the reduced gradient. SIAM Journal on Control and Optimization 14/1, 42-61.

J.Ch. Gilbert (1986,a). Une méthode à métrique variable réduite en optimisation avec contraintes d'égalité non linéaires. Rapport de recherche de l'INRIA RR-482. 78153 Le Chesnay, France. (to appear, in part, in Mathematical Modelling and Numerical Analysis).

J.Ch. Gilbert (1986,b). Une méthode de quasi-Newton réduite en optimisation sous contraintes avec priorité à la restauration. Lecture Notes in Control and Information Sciences 83, 40-53. A. Bensoussan, J.L. Lions, (eds.). Springer-Verlag.

J.Ch. Gilbert (1986,c). Sur quelques problèmes d'identification et d'optimisation rencontrés en physique des plasmas. Thèse de doctorat de l'université Paris VI.

P.E. Gill, W. Murray, M.A. Saunders, G.W. Stewart, M.H. Wright (1985). Properties of a representation of a basis for the null space. Mathematical Programming 33, 172-186.

S.P. Han (1976). Superlinearly convergent variable metric algorithms for general nonlinear programming problems. Mathematical Programming 11, 263-282.

S.P. Han (1977). A globally convergent method for nonlinear programming. Journal of Optimization Theory and its Application 22/3, 297-309.

N. Maratos (1978). Exact penalty function algorithms for finite dimensional and control optimization problems. Ph.D. Thesis, University of London.

D.Q. Mayne, E. Polak (1982). A superlinearly convergent algorithm for constrained optimization problems. Mathematical Programming Study 16, 45-61.

H. Mukai, E. Polak (1978). On the use of approximations in algorithms for optimization problems with equality and inequality constraints. SIAM Journal on Numerical Analysis 15/4, 674-693.

J. Nocedal, M.L. Overton (1985). Projected hessian updating algorithms for nonlinearly constrained optimization. SIAM Journal on Numerical Analysis 22/5, 821-850.

M.J.D. Powell (1978). The convergence of the variable metric methods for nonlinearly constrained optimization calculations. Nonlinear Programming 3, O.L. Mangasarian, R.R Meyer, S.M. Robinson, eds., Academic Press, 27-63.

J.B. Rosen (1961). The gradient projection method for nonlinear programming. Part II: nonlinear constraints. Journal of the Society for Industrial and Applied Mathematics (SIAM) 9/4, 514-532.

J. Stoer (1984). Principles of sequential quadratic programming methods for solving nonlinear programs. Proceedings of the NATO ASI on Computational Mathematical Programming, Bad Windsheim, Germany.

R.B. Wilson (1963). A simplicial algorithm for concave programming. Ph.D. thesis. Graduate School of Business Administration, Harvard Univ., Cambridge, MA.

Y. Yuan (1985). An only 2-step Q-superlinear convergence example for some algorithms that use reduced hessian approximations. Mathematical Programming 32, 224-231.

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique